



Beker, W. L., Van der Kamp, M. W., Mulholland, A. J., & Sokalski, W. A. (2017). Rapid estimation of catalytic efficiency by cumulative atomic multipole moments: application to ketosteroid isomerase mutants. *Journal of Chemical Theory and Computation*, 13(2), 945-955. <https://doi.org/10.1021/acs.jctc.6b01131>

Peer reviewed version

License (if available):
CC BY-NC

Link to published version (if available):
[10.1021/acs.jctc.6b01131](https://doi.org/10.1021/acs.jctc.6b01131)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via ACS at <http://pubs.acs.org/doi/abs/10.1021/acs.jctc.6b01131>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

SUPPORTING INFORMATION

Rapid estimation of catalytic efficiency by cumulative atomic multipole moments: application to ketosteroid isomerase mutants

Wiktor Beker¹, Marc van der Kamp^{2,3,4*}, Adrian Mulholland^{2,3}, W. Andrzej Sokalski^{*1}

¹Advanced Materials Engineering and Modelling Group, Faculty of Chemistry, Wrocław University of Technology, Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland

²School of Biochemistry, Biomedical Sciences Building, University Walk, Bristol, BS8 1TD, UK

³BrisSynBio Synthetic Biology Research Centre, Life Sciences Building, Tyndall Avenue, University of Bristol, BS8 1TQ, UK

⁴Centre of Computational Chemistry, School of Chemistry, Cantock's Close, University of Bristol, BS8 1TS, UK.

Contents

Molecular dynamics simulations of KSI variants complexed with the intermediate state (IS).....	3
Convergence of Differential Intermediate State Stabilization (DISS) energy calculated by Cumulative Atomic Multipole Moments (CAMM).....	6
Linear regression parameters	9
Comparison of DISS obtained with CAMM or with fixed point charges	10
Comparison with Chakravorty and Hammes-Schiffer (Biochemistry, 2009)	12
Testing the importance of many-body polarization	13
The influence of structure on SAPT results and mutual cancelling of interaction energy terms.....	14
The effect of substrate truncation	16
Comparison of computational requirements of CAMM, HF and SAPT0 calculations	16
REFERENCES	17

Molecular dynamics simulations of KSI variants complexed with the intermediate state (IS)

System set-up and parameters

The protein coordinates of the KSI dimer were taken from PDB entry 1OHP, chains A and B, with Asn38 mutated back to the wild-type catalytic base Asp38. In chain A, the intermediate of the KSI reaction was included, based on the coordinates of the co-crystallised inhibitor 5 α -estran-3,17-dione, whereas 5 α -estran-3,17-dione was removed from the active site of chain B. Asp99 was treated as protonated in both chains (in line with its role in the enzyme and predicted pK_a values from PropKa3.1¹ of 8.2 in chain A with the substrate bound and 9.2 in chain B with an empty active site), and Asp38 was treated as protonated only in chain A, based on the mechanism (Asp38 has abstracted the proton from C4 of the substrate to form the intermediate; see scheme in main manuscript). All other ionisable residues were treated in their standard protonation states (Asp/Glu negatively charged, Arg/Lys positively charged, His neutral). All three histidines in both chains were modeled as the tautomer protonated on N ϵ 2 and the functional groups of the following residues were ‘flipped’ (rotated by 180°) as predicted by the Optimal Hydrogen Bonding Network²: Asn2, Asn104, His100 and His122 in chain A, and Asn19, Asn57 and Asn104 in chain B.

All crystallographic waters were kept and additional water molecules were added using the *solvate* plugin to VMD³ (with a ‘buffer’ for overlapping water deletion of 2.3) such that the edge of the waterbox was at least 11 Å away from any protein atom. The *autoionize* plugin was then used to replace 6 bulk waters with Na⁺ to neutralise the system. The resulting PSF file (generated by the VMD plugin *psfgen*) was converted by the AmberTools14 utility *chamber* into a prmtop file that can be used by AMBER. This ensures that CHARMM force-fields are accurately interpreted by AMBER⁴. The CHARMM36⁵ force-field was used for the protein, TIP3P⁶ for water and CgenFF⁷ parameters (atom types assigned using <https://cgenff.paramchem.org>) with CHELPG charges (obtained with the RED Server⁸) for the intermediate state. The resulting topology file is given below.

CgenFF topology file for the KSI substrate in the intermediate state, with CHELPG charges obtained through the RED Server⁸:

```
AUTO ANGLES DIHE
RESI IS1          -1.000 !
GROUP            ! CHELPG charges
ATOM C1          CG321    -0.0333
ATOM H1A         HGA2     -0.0210
ATOM H1B         HGA2     -0.0210
ATOM C2          CG321    -0.1135
ATOM H2A         HGA2     0.0105
ATOM H2B         HGA2     0.0105
ATOM C3          CG2D1O   0.6927
ATOM O3          OG312    -0.8009
ATOM C4          CG2DC1   -0.8064
ATOM H4B         HGA4     0.1680
ATOM C5          CG2DC2   0.2701
ATOM C6          CG2DC2  -0.6545
```

ATOM	H6A	HGA4	0.1585
ATOM	C7	CG321	0.2989
ATOM	H7A	HGA2	-0.0719
ATOM	H7B	HGA2	-0.0719
ATOM	C8	CG311	0.0312
ATOM	H8A	HGA1	-0.0308
ATOM	C9	CG311	0.0218
ATOM	H9A	HGA1	-0.0559
ATOM	C10	CG301	0.2475
ATOM	C19	CG331	-0.1532
ATOM	H13A	HGA3	0.0118
ATOM	H13B	HGA3	0.0118
ATOM	H13C	HGA3	0.0118
ATOM	C11	CG321	0.1238
ATOM	H11A	HGA2	-0.0405
ATOM	H11B	HGA2	-0.0405
ATOM	C12	CG321	-0.1188
ATOM	H12A	HGA2	0.0000
ATOM	H12B	HGA2	0.0000
ATOM	C13	CG3RC1	0.2651
ATOM	C18	CG331	-0.2838
ATOM	H18A	HGA3	0.0576
ATOM	H18B	HGA3	0.0576
ATOM	H18C	HGA3	0.0576
ATOM	C14	CG3RC1	-0.0088
ATOM	H14A	HGA1	-0.0113
ATOM	C15	CG3C52	-0.0483
ATOM	H15A	HGA2	0.0043
ATOM	H15B	HGA2	0.0043
ATOM	C16	CG3C52	-0.1361
ATOM	H16A	HGA2	0.0431
ATOM	H16B	HGA2	0.0431
ATOM	C17	CG205	0.5131
ATOM	O17	OG2D3	-0.5923
BOND	C1	H1A	
BOND	C1	H1B	
BOND	C1	C2	
BOND	C1	C10	
BOND	C2	H2A	
BOND	C2	H2B	
BOND	C2	C3	
BOND	C3	O3	
BOND	C3	C4	
BOND	C4	H4B	
BOND	C4	C5	
BOND	C5	C6	
BOND	C5	C10	
BOND	C6	H6A	
BOND	C6	C7	
BOND	C7	H7A	
BOND	C7	H7B	
BOND	C7	C8	

```

BOND C8      H8A
BOND C8      C9
BOND C8      C14
BOND C9      H9A
BOND C9      C10
BOND C9      C11
BOND C10     C19
BOND C19     H13A
BOND C19     H13B
BOND C19     H13C
BOND C11     H11A
BOND C11     H11B
BOND C11     C12
BOND C12     H12A
BOND C12     H12B
BOND C12     C13
BOND C13     C18
BOND C13     C14
BOND C13     C17
BOND C18     H18A
BOND C18     H18B
BOND C18     H18C
BOND C14     H14A
BOND C14     C15
BOND C15     H15A
BOND C15     H15B
BOND C15     C16
BOND C16     H16A
BOND C16     H16B
BOND C16     C17
BOND C17     O17
IMPR C17      C16      C13      O17

```

Simulation and analysis - methods

All simulations were performed with the AMBER package v. 14. Two independent simulations were performed for each mutant (with randomized positions of the Na⁺ ions prior to initial minimisation).

Minimisations were performed with sander (on single CPUs) and MD simulations with pmemd.MPI (on 8 CPUs). Each system was prepared for MD simulation by initial minimisation of all solvent, the substrate and all hydrogens (300 steps), a brief (50 ps) MD simulation of all solvent and finally a minimisation of all atoms (300 steps) with positional restraints on the protein C α atoms of 5.0 kcal mol⁻¹ Å⁻². The system was then heated (after random assignment of velocities at 25 K) to 298 K in 20 ps, using Langevin dynamics for temperature control (collision frequency of 1 ps), again with the positional restraints on the protein C α atoms. The positional restraints were gradually released and the pressure equilibrated to 1 atm in 4 steps of 10 ps, using the Berendsen barostat (pressure relaxation time 1 ps). Finally, 1 ns simulations were performed in the NPT ensemble at 298 K and 1 atm (with Langevin dynamics for temperature control and the Berendsen barostat). Throughout all MD simulations, a one-sided restraint was applied to ensure that the distance between Asp38 O δ 2 and C4 of the substrate does not become larger than 2.65 Å, so that the conformations sampled are in line with the reaction mechanism. Further, a 2

fs timestep was used with SHAKE applied to bonds involving hydrogen, the cut-off for direct-space non-bonded interactions was 8 Å and long-range electrostatics were treated using the Particle Mesh Ewald method.

The 2 trajectories obtained from two 1 ns simulations (saved every 1 ps) were clustered based on the mass-weighted RMSD of the substrate, Asp38 (the catalytic base) and the residues at positions 14 & 99 in chain A. The hierarchical agglomerative algorithm was used and 4 clusters were requested, using the AmberTools program cpptraj⁹.

Convergence of Differential Intermediate State Stabilization (DISS) energy calculated by Cumulative Atomic Multipole Moments (Camm)

To test whether the multipolar expansion of the differential intermediate state stabilization (DISS) in KSI converges, calculations were performed at different truncation levels L ($L=L_a+L_b$, the sum of ranks of multipole moments of monomer a and b , related to the inverse distance term in the multipole expansion via $R^{-(L+1)}$ (exponent truncation), see ref.¹⁰). Calculations were performed for all KSI variants considered (Figures S1-S5), employing the whole (rather than the truncated) substrate (see Figure 3) and structures obtained with the minimal perturbation approach (see main text). The Camm expansion of the DISS energy is close to convergence after the R^{-4} term ($L=3$) in most cases. Results diverge after R^{-6} or R^{-7} ($L=5$ or 6, respectively) in some cases, which is likely related to numerical errors in the determination of high-rank; see also further discussion in the main text. Overall, the truncation level R^{-5} ($L=4$) is a reasonable choice.

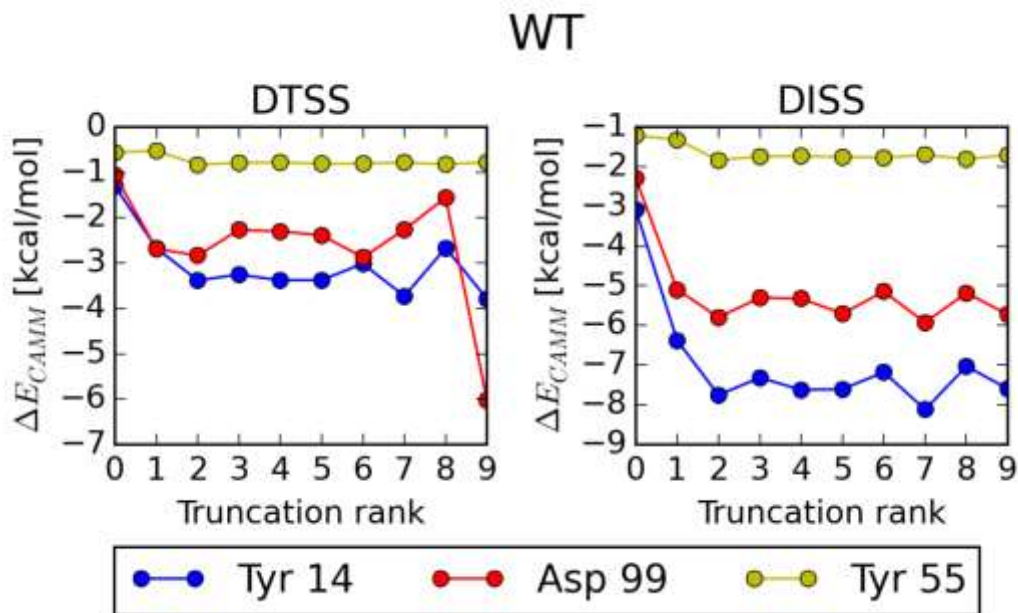


Figure S1. Convergence of Camm DISS and DTSS energy in wild-type KSI (QM/MM structure from ref.¹¹). The multipolar contribution to the differential intermediate state stabilization as calculated with various truncation levels of Camm expansion.

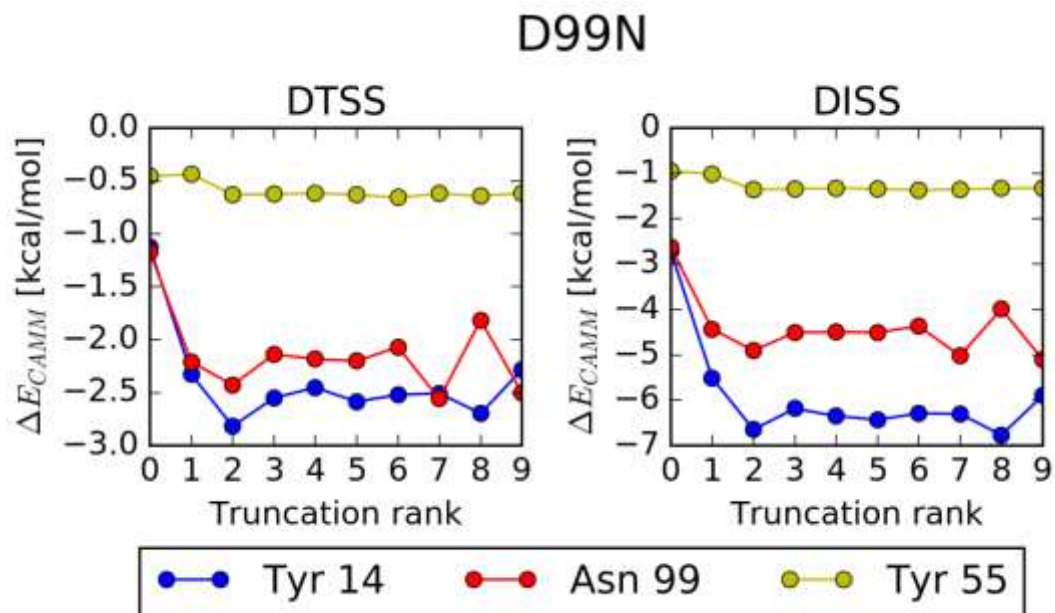


Figure S2. Convergence of CAMM DISS and DTSS energy in the D99N KSI mutant. The multipolar contribution to differential intermediate state stabilization calculated with various truncation levels of CAMM expansion.

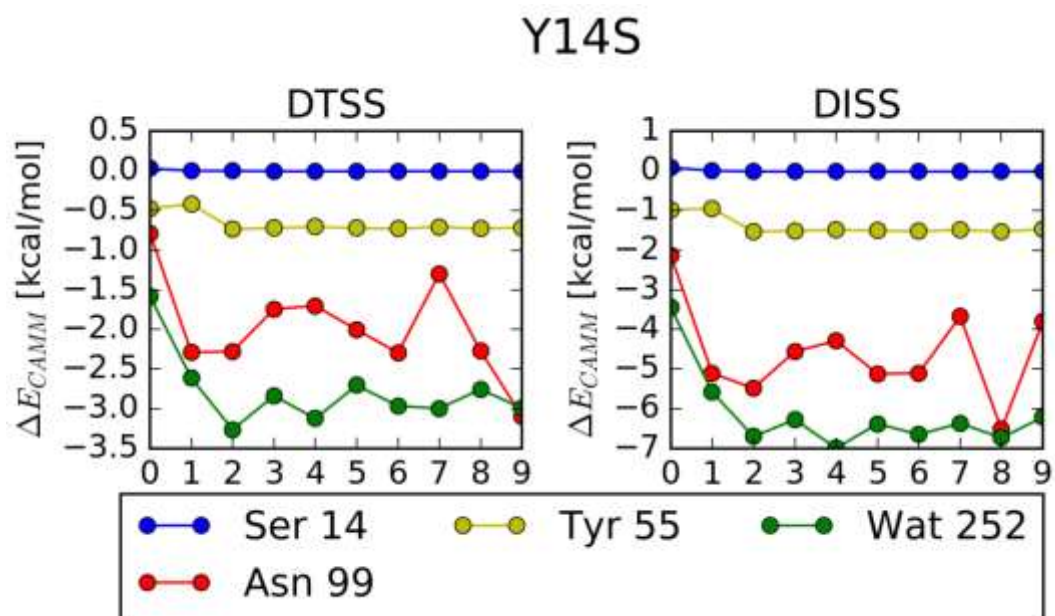


Figure S3. Convergence of CAMM DISS and DTSS energy in the Y14S KSI mutant. The multipolar contribution to differential intermediate state stabilization calculated with various truncation levels of CAMM expansion.

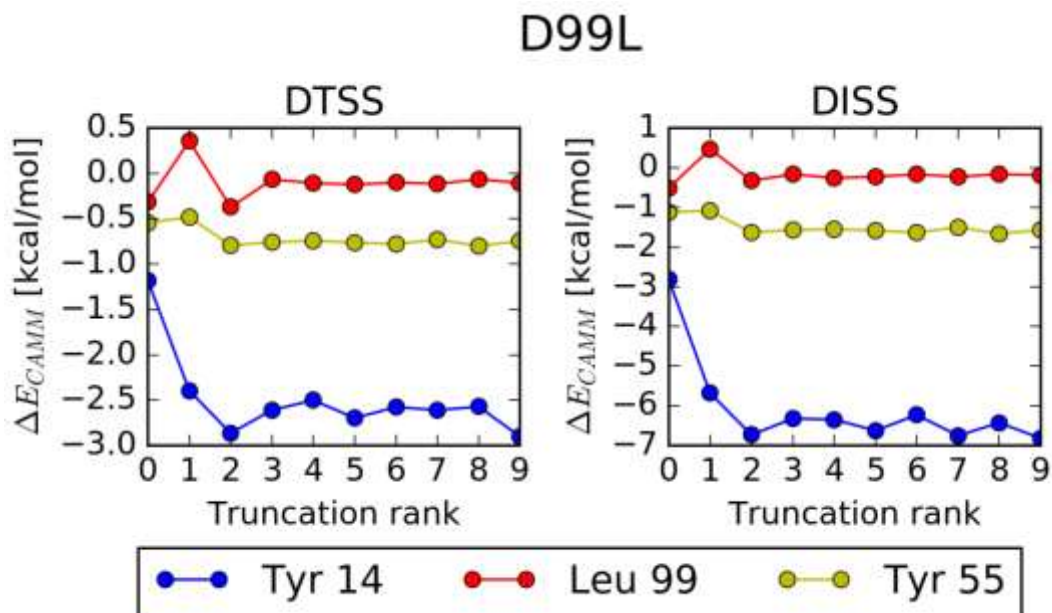


Figure S4. Convergence of CAMM DISS and DTSS energy in the D99L KSI mutant. The multipolar contribution to differential intermediate state stabilization calculated with various truncation levels of CAMM expansion.

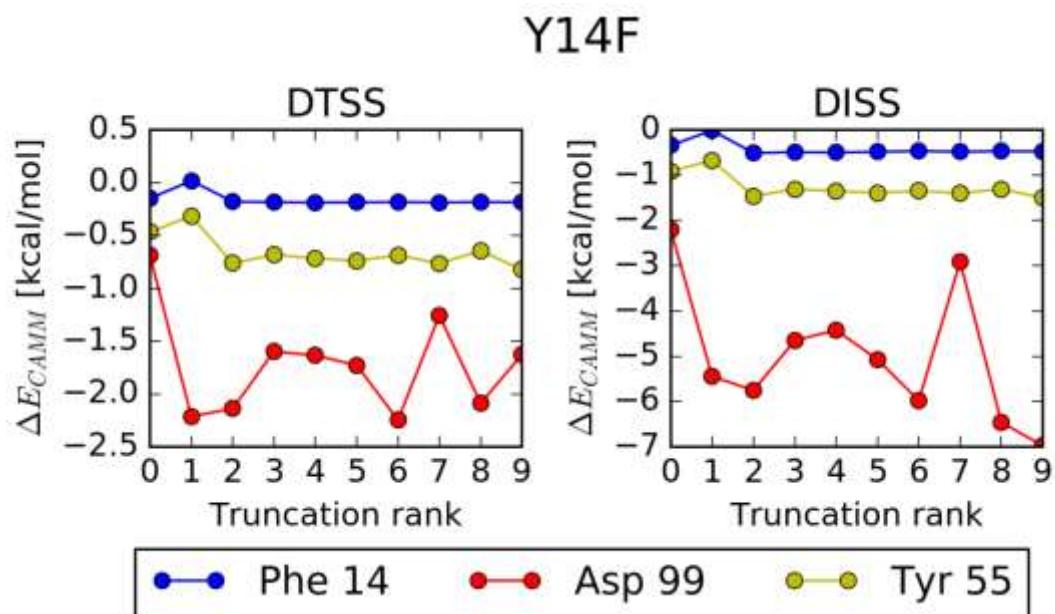


Figure S5. Convergence of CAMM DISS and DTSS energy in the Y14F KSI mutant. The multipolar contribution to differential intermediate state stabilization calculated with various truncation levels of CAMM expansion.

Linear regression parameters

Table S1. Best-fit parameters, Pearson correlation coefficients (ρ_P) and Spearmann rank correlation coefficients (ρ_S) between differential intermediate state stabilization (DISS) and experimental apparent free energy barriers (calculated from the works of Kraut and Choi^{12,13}, see Computational details). Calculations are based on the cluster centroid structures obtained from MD simulation (see Computational details).

MD clusters

Method	All mutants				Without Y14S			
	slope	intercept	ρ_P	ρ_S	slope	intercept	ρ_P	ρ_S
$\Delta E_{MTP}^{(10)}(CAMM)$	0.47 ± 0.12	17.9 ± 1.2	0.92	0.9	0.53 ± 0.08	18.21 ± 0.71	0.98	1.0
$\Delta E_{elst}^{(10)}$	0.37 ± 0.09	18.0 ± 1.2	0.93	0.9	0.41 ± 0.08	18.21 ± 0.9	0.97	1.0
ΔE_{HF}	0.50 ± 0.12	20.3 ± 1.7	0.93	1.0	0.56 ± 0.09	20.7 ± 1.2	0.98	1.0
ΔE_{SAPT0}	0.5 ± 0.1	20.2 ± 1.5	0.94	1.0	0.51 ± 0.08	20.5 ± 1.2	0.98	1.0

Table S2. Best-fit parameters, Pearson correlation coefficients (ρ_P) and Spearmann rank correlation coefficients (ρ_S) between differential intermediate state stabilization (DISS) and experimental apparent free energy barriers (calculated from the works of Kraut and Choi^{12,13}, see Computational details). Calculations are based on the ‘minimal perturbation’ structures (see Computational details).

Minimal Perturbation Structures

Method	All mutants				Without Y14S			
	slope	intercept	ρ_P	ρ_S	slope	intercept	ρ_P	ρ_S
$\Delta E_{MTP}^{(10)}(CAMM)$	0.52 ± 0.12	18.8 ± 1.2	0.94	1.0	0.57 ± 0.06	19.0 ± 0.6	0.99	1.0
$\Delta E_{elst}^{(10)}$	0.32 ± 0.15	18.5 ± 2.4	0.78	0.3	0.32 ± 0.19	18.0 ± 3.0	0.78	0.2
ΔE_{HF}	0.4 ± 0.1	19.3 ± 1.4	0.94	1.0	0.45 ± 0.12	19.3 ± 1.6	0.94	1.0
ΔE_{SAPT0}	0.48 ± 0.06	20.4 ± 0.9	0.98	1.0	0.48 ± 0.07	20.4 ± 1.1	0.98	1.0

Comparison of DISS obtained with CAMM or with fixed point charges

Here, we compare the use of CAMM with the use of a fixed point charge model for calculating Differential Intermediate State Stabilization for KSI variants. Point charges were CHELPG charges as obtained with the RED Server ⁸ for the reactant and intermediate states. Both approaches have a similar trend, but it should be noted that the point-charge approach benefits from a significant cancellation of errors (Table S3); Note that DISS (defined as $E_{\text{int}}(\text{IS}) - E_{\text{int}}(\text{RS})$) is dominated by the interaction energy with the intermediate state, which in turn is quite well described by a point charge model (due to its non-zero charge). This is reflected in the fact that the difference between the point charge (FF) and CAMM interaction energies is small this state (compared to the reactant state). The errors in interaction energies of the point charge model (compared to the CAMM model) are not systematic, indicating that changing the point charges used cannot easily improve the results.

Table S3. Comparison of CAMM and FF charges energies ('Minimal perturbation' structure)

MUTANT	CAMM RS [kcal/mol]	FF RS [kcal/mol]	δ RS [kcal/mol]	CAMM IS [kcal/mol]	FF IS [kcal/mol]	δ IS [kcal/mol]	CAMM DISS [kcal/mol]	FF DISS [kcal/mol]	δ DISS [kcal/mol]
WT	-23.15	-18.95	4.20	-36.24	-34.35	1.89	-13.09	-15.40	-2.31
D99N	-20.55	-16.11	4.44	-32.87	-30.92	1.95	-12.32	-14.80	-2.48
Y14S	-22.28	-18.07	4.21	-34.07	-31.50	2.57	-11.79	-13.44	-1.64
D99L	-14.48	-12.18	2.30	-22.42	-21.11	1.31	-7.94	-8.93	-0.99
Y14F	-13.29	-11.10	2.18	-19.03	-18.58	0.45	-5.74	-7.47	-1.73
<i>R</i>	0.86	0.85		0.90	0.91		0.94	0.94	
<i>R</i> _(noY14S)	0.96	0.96		0.98	0.97		0.99	0.97	

For both models (CAMM and force field point charges FF), interaction energies with substrate (RS) and intermediate (IS) are shown, as well as their difference (Differential Intermediate Stabilization, DISS) and the deviation of the point charge model from CAMM (δ s) for all the values. For interaction energies, Pearson correlation coefficients (*R*) with experimental apparent free energy barriers (calculated from the works of Kraut and Choi^{12,13}, see Computational details) are given, with and without inclusion of Y14S (*R*_(no Y14S)).

Further comparison of CAMM with point charge approximations is provided in Figure S6. Here, we consider reactants and Asp99; by increasing the separation between their centers of mass, we monitor Asp99 contribution to the DISS as a function of distance difference DX (with respect to the value from QM/MM study). As a reference we took the DMA multipole moments¹⁴. The results indicate that point charge models cannot reproduce charge anisotropy well in comparison to CAMM.

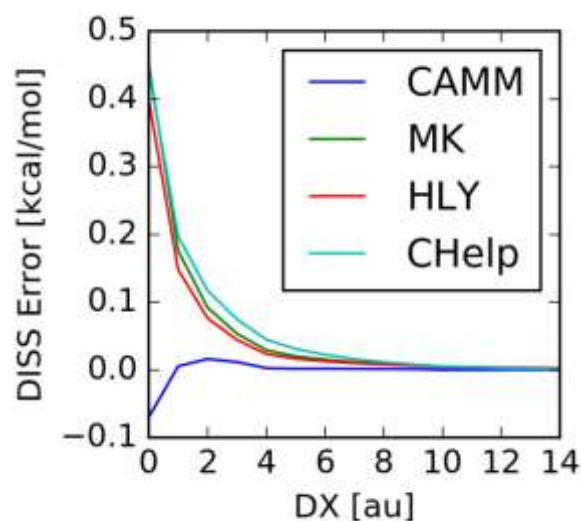


Figure S6. Error in DISS contribution of Asp 99 as a function of the separation with respect to the value from the QM/MM study. DMA multipoles serve as a reference. CAMM- Cumulative Atomic multipole moments (up to R^{-5}); MK-Merz-Singh-Kollman^{15,16}, HLY-Hu,Lu,Yang¹⁷, CHelpG¹⁸. DX- distance by which the separation of centers of mass is varied; zero value corresponds to equilibrium geometry.

Comparison with Chakravorty and Hammes-Schiffer (Biochemistry, 2009)

Table S4. Comparison of intermediate stabilization energies (all in kcal/mol) between Chakravorty and Hammes-Schiffer's work¹⁹ and the present study. Apparent activation barriers calculated basing on experimental data from Kraut and Choi^{12,13} (see *Computational details*).

Enzyme	ΔG_{app}^\ddagger ^{12,13}	$E_{EL}(RS)$ ¹⁹	$E_{EL}(IS)$ ¹⁹	$E_{EL}(IS)-E_{EL}(RS)$	DISS(CAMM)
WT	11.5	-30	-101	-71	-15
D99L	14.2	-24	-84	-60	-8
Y14F	16.0	-27	-91	-64	-6

The differences between Chakravorty and Hammes-Schiffer's work¹⁹ and our work arise mainly from different sizes of system used in both set of calculations (in our work, the Asp38 side-chain was included in the system considered for the interaction energy calculations, in contrast to Chakravorty and Hammes-Schiffer's work). However, applying our approach to point-charge model of whole proteins leads to values consistent with presented CAMM results and of magnitudes comparable to those from EVB study¹⁹ (see Table S5). There are also minor technical differences impeding direct numerical comparison of these two approaches. Firstly, different PDB structures were used as a starting point for modeling (PDB code 1QJG in EVB study and 1OHP here).

Table S5. Interaction energies (in kcal/mol) between androstedione and its environment calculated with point charges (structures obtained from clustering of MD simulations). The values are similar to those presented in Table S4 in the work of Chakravorty and Hammes-Schiffer¹⁹.

Protein	$E_{EL}(RS)$	$E_{EL}(IS)$	$E_{EL}(IS)-E_{EL}(RS)$
WT	-42	-126	-84
D99N	-45	-123	-78
Y14S	-46	-133	-87
D99L	-40	-120	-80
Y14F	-42	-126	-84

Testing the importance of many-body polarization

To examine the three-body effects associated with polarization, we chose the Tyr14-Tyr55 interaction as a test-case. The hydrogen bond donated by Tyr55 to the Tyr14 –OH may polarize this group and thus strengthen its O-H dipole. To include this effect, the wavefunction was calculated for both residues treated as one system and used for the determination of CAMMs. Subsequently, DISS calculations with this polarized Tyr14 system were compared with DISS calculations using including the individual, non-polarized tyrosine residues (Table S6). In case of the Y14S and Y14F mutants, the same strategy was applied, including Ser14, W_{TyrOH} and Tyr55 in the former, and Phe14 and Tyr55 in the latter case. All three variants that contain the Tyr14-Tyr55 system show a similar influence of polarization: the differential stabilization decreases (i.e. stronger relative stabilization of the intermediate state) by about 0.5 kcal/mol compared to the non-polarized model. Thus, the error arising from the neglecting Tyr14-Tyr55 polarization is ~7% of the ‘polarized’ value. It is further found that the Tyr55 contribution is nearly constant among mutants (~ –1.5 kcal/mol) and we expect similar behavior from other residues common in all considered variants.

In Y14F, the polarization effect of Tyr55 *increases* the DISS (i.e. weaker relative stabilization of the intermediate state compared to the reactant state); it’s worth noting that in this mutant, the value presented in Table S6 is dominated by Tyr55 (–1.35 kcal/mol). On the other hand, in Y14S the difference is large in comparison with other variants (about 1.2 kcal/mol). This mutation is more difficult to model using the DISS(CAMM) approach (see discussion in main text).

Table S6. Comparison of DISS energies (in kcal/mol) calculated using a non-polarized (per-residue) CAMM model and one that includes explicit residue 55 – residue 14 polarization.

Mutant	Residues polarized	Residues unpolarized	Difference (pol.-unpol.)
WT	-10.02	-9.35	-0.67
D99N	-8.10	-7.68	-0.42
Y14S	-9.70	-8.51	-1.20
D99L	-8.40	-7.89	-0.50
Y14F	-1.63	-1.85	0.23

The influence of structure on SAPT results and mutual cancelling of interaction energy terms

In the main text, it was shown that the difference between DISS(CAMM) energies obtained for Minimal Perturbation (MP) structures and MD cluster centroids (MD) is small or at least nearly systematic. Here we present a similar comparison for other terms as defined by SAPT (Figure S7). The most dramatic case is the D99L mutant, where the short-range penetration term is responsible for overestimation of Leu99 electrostatic ($\Delta E_{EL}^{(10)}$) contribution. Other levels (ΔE_{HF} and ΔE_{SAPT0}) are less affected by the structural change by mutual cancelling of those short-range terms (see Figure S8). However, this change is larger and less systematic than that observed in CAMM (see main text).

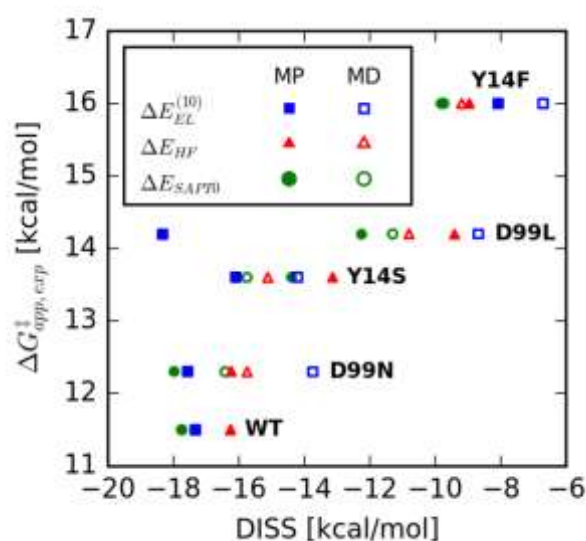


Figure S7. DISS energies calculated with different levels in SAPT decomposition, using ‘minimal perturbation’ (MP) structures and MD trajectory cluster centroids (MD).

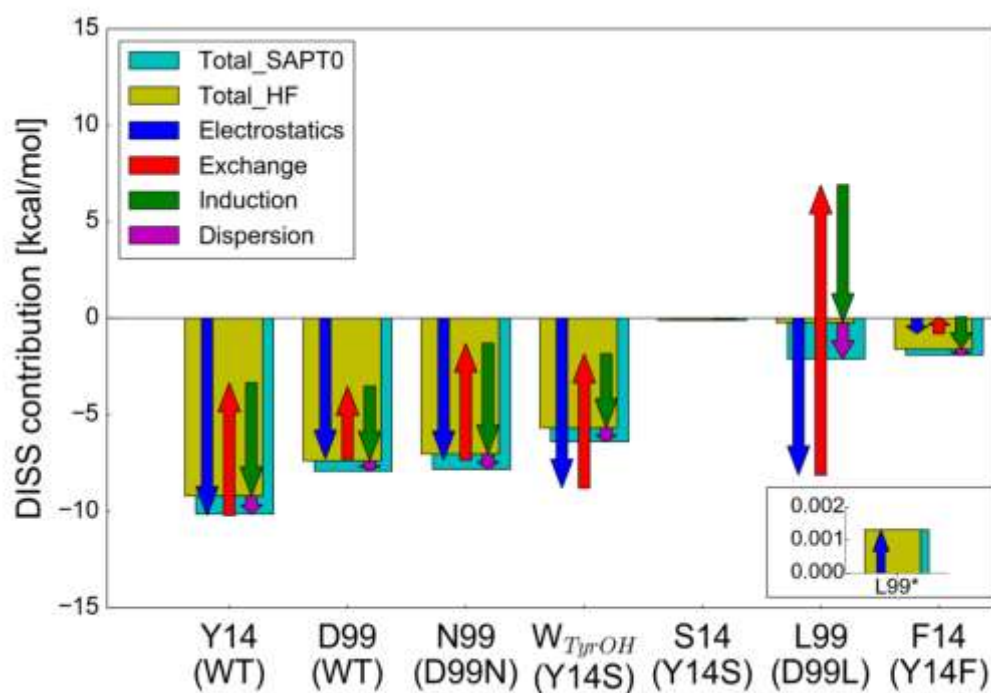


Figure S8. A decomposition of DISS contribution arising from analyzed amino acids (in Minimal Perturbation geometries). Subsequent contributions (according to SAPT²⁰ decomposition) are marked as arrows and aligned head-to-tail.

The effect of substrate truncation

We provide details of comparison between results obtained with full substrate and the one truncated to the two rings encompassing atoms participating in reaction (see Computational Details in main text). As can be seen in Table S7, the maximum error is smaller than 0.5 kcal/mol for overall DISS, with errors for DISS calculated with CAMM are the lowest (<0.25 kcal/mol).

Table S7. DISS error arising from truncation of the substrate ($\text{DISS}_{\text{whole}} - \text{DISS}_{\text{truncated}}$, kcal/mol).

	Tyr 14	Asp 99	Tyr 55	Sum
$\Delta E_{MTP}^{(10)}(\text{CAMM})$	0.06	0.09	0.10	0.24
$\Delta E_{elst}^{(10)}$	0.06	0.13	0.09	0.27
ΔE_{HF}	0.16	0.15	0.11	0.43
ΔE_{SAPT0}	0.15	0.15	0.11	0.41

Comparison of computational requirements of CAMM, HF and SAPT0 calculations

To indicate the difference in computational effort of CAMM, HF(SCF) and SAPT0 calculations, we compare the computer times required for evaluation of electrostatic interaction energies for homodimers within S22 training set²¹, see Figure S9. (Hereafter, A and N state for number of atoms and basis set functions in monomers, respectively.) Note that CAMM(R^{-5}) calculations for this training set take only a few milliseconds, scaling approximately as $O(A^2)$, a negligible value compared to the SAPT0 electrostatic term evaluation in the dimer basis set, scaling as $O((2N)^4)$. It should be mentioned that CAMM energy calculations of any new chemical species requires two SCF runs for each monomer, each one scaling as $O(N^4)$, to determine multipole moments (see the SCF series in Fig. S9). However, with a library of pre-computed CAMMs for aminoacids and one initial SCF calculation for each ligand/transition state, evaluation of interaction energies with any number of amino acid residues (including e.g. different enzyme variants) will follow the CAMM curve rather than the SCF curve presented in Fig. S9.

All times presented in Figure S9 were evaluated using the same machine, containing two Intel® Xeon® E5-2630 v2 @ 2.60 GHz CPUs. SCF for monomers from S22 training set²¹ were obtained using GAMESS²² with 12 threads, whereas SAPT0 energies were calculated with PSI4²³, using the same number of threads. All computations were done with aug-cc-pVDZ basis set.

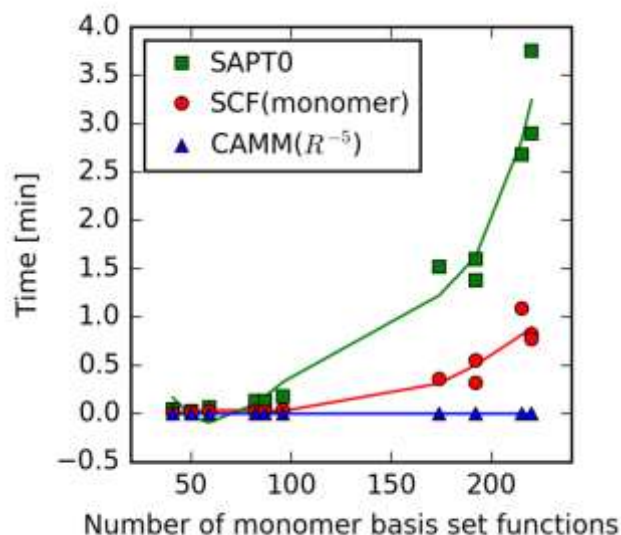


Figure S9. Comparison of computational times of evaluation of CAMM and SAPT0 energies in homodimers from S22 training set²¹. System size is measured as total number of basis set functions in dimer. Here, SCF series mean total time required to determine CAMMs for each monomer. Note that sole CAMM interaction energy calculation takes a few milliseconds.

REFERENCES

- (1) Sondergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, 7 (7), 2284–2295.
- (2) Hooft, R. W. W.; Sander, C.; Vriend, G. *Proteins* **1996**, 26, 363–376.
- (3) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph.* **1996**, 14, 33–38.
- (4) Crowley, M. F.; Williamson, M. J.; Walker, R. C. *Int. J. Quantum Chem.*, **2009**, 109, 3767–3772.
- (5) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell Jr., A. *D. J. Chem. Theory Comput.* **2012**, No. 8, 3257–3273.
- (6) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, 79, 926.
- (7) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; MacKerell Jr., A. D. *J. Comput.*

- Chem.* **2010**, *31*, 671–690.
- (8) Vanquelef, E.; Simon, S.; Marquant, G.; Garcia, E.; Klimerak, G.; Delepine, J. C.; Cieplak, P.; Dupradeau, F. Y. *Nucleic Acids Res.* **2011**, *39*, W511–W517.
- (9) Roe, D. R.; Cheatham, T. E. I. *J. Chem. Theory Comput.* **2013**, *9* (7), 3084–3095.
- (10) Sokalski, W. A.; Poirier, R. A. *Chem. Phys. Lett.* **1983**, *98* (1), 86–92.
- (11) van der Kamp, M. W.; Chaudret, R.; Mulholland, A. J. *Febs J.* **2013**, *280* (13), 3120–3131.
- (12) Kraut, D. A.; Sigala, P. A.; Fenn, T. D.; Herschlag, D. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (5), 1960–1965.
- (13) Choi, G.; Ha, N. C.; Kim, M. S.; Hong, B. H.; Oh, B. H.; Choi, K. Y. *Biochemistry* **2001**, *40* (23), 6828–6835.
- (14) Stone, A. J. *Chem. Phys. Lett.* **1981**, *83* (2), 233.
- (15) Besler, B. H.; Merz, K. M. J.; Kollman, P. A. *J. Comput. Chem.* **1990**, *11*, 431–439.
- (16) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129–145.
- (17) Hu, H.; Lu, Z.; Yang, W. *J. Comput. Chem.* **2007**, *3*, 1004–1013.
- (18) Chirlian, L. E.; Francl, M. M. *J. Comput. Chem.* **1987**, *8*, 894–905.
- (19) Chakravorty, D. K.; Soudackov, A. V; Hammes-Schiffer, S. *Biochemistry* **2009**, *48* (44), 10608–10619.
- (20) Cwiok, T.; Jeziorski, B.; Kolos, W.; Moszynski, R.; Szalewicz, K. *Theochem-Journal Mol. Struct.* **1994**, *113*, 135–151.
- (21) Jurecka, P.; Sponer, J.; Cerny, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8* (17), 1985–1993.
- (22) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14* (11), 1347–1363.

- (23) Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F. A.; Fermann, J. T.; Mintz, B. J.; Burns, L. A.; Wilke, J. J.; Abrams, M. L.; Russ, N. J.; Leininger, M. L.; Janssen, C. L.; Seidl, E. T.; Allen, W. D.; Schaefer, H. F.; King, R. A.; Valeev, E. F.; Sherrill, C. D.; Crawford, T. D. *Wiley Interdiscip. Rev. Mol. Sci.* **2012**, 2 (4), 556–565.